

## Tilburg University

### **Can we identify the poorest quality of life? Assessing the importance of quality of life using the WHOQOL-100**

Skevington, S.M.; O'Connell, K.A.; WHOQOL, G.; van Heck, G.L.; de Vries, J.

*Published in:*  
Quality of Life Research

*Publication date:*  
2004

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*  
Skevington, S. M., O'Connell, K. A., WHOQOL, G., van Heck, G. L., & de Vries, J. (2004). Can we identify the poorest quality of life? Assessing the importance of quality of life using the WHOQOL-100. *Quality of Life Research*, 13, 23-34.

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Can we identify the poorest quality of life? Assessing the importance of quality of life using the WHOQOL-100<sup>☆</sup>

Suzanne M. Skevington, Kathryn A. O'Connell & the WHOQOL Group  
*WHO Centre for the Study of Quality of Life, Department of Psychology, University of Bath,  
Bath BA2 7AY, UK (E-mail: s.m.skevington@bath.ac.uk)*

Accepted in revised form 6 February 2003

### Abstract

In this study, WHOQOL survey data obtained from 4802 sick and well participants in 15 countries were used to investigate the relationship between judgements about different dimensions of quality of life (QOL) (core scores) and the importance attributed to them. As a theoretical framework, we applied the WHOQOL Group's (1995) definition of QOL which indicates that those who report the very poorest QOL will be least likely to have met their own '...goals, expectations, standards and concerns'. Those with the poorest QOL would therefore be expected to show the biggest difference between core and importance scores, and therefore be distinguishable from respondents whose QOL was poor, better or best. The main effects from overall analyses confirmed that those reporting the largest negative differences tended to report the poorest QOL and also attached a high degree of importance to these dimensions. Evidence for a decreasing differential across the four groups (poorest to best) was confirmed for the majority (18) of facets. However facet level analyses comparing groups with different levels of QOL showed that only five facets distinguished those with the poorest QOL from those whose QOL was poor, so the theory is not well supported. Furthermore the contribution of core-importance facet differences reduced the overall prediction of QOL, when compared with a regression of core scores alone. Importance information about specific facets may have limited potential to be used alongside the main instrument to identify areas of the poorest QOL for clinical or social action.

**Key words:** Assessment, Health, Importance, Quality of life, WHOQOL

### Introduction

Quality of life (QOL) in health care is frequently conceptualised as a multi-dimensional phenomenon [1–4]. But how important are the various dimensions that are included in multi-dimensional profiles? Without evaluating their importance, the assumption embedded in some instruments is that all aspects of QOL are equally important at every cultural or social group level. Furthermore that individuals perceive all aspects of their QOL to be equally important, and this is a shortcoming in

their design. Kind observes 'Most QOL assessments suffer from the limitations that undermine their potential – notably the absence of any individual preferences indicating the importance to be attached to different aspects of health status and QOL indicators' [5]. Information about importance is useful because it can assist instrument developers in selecting dimensions most relevant to the population for whom the scale is intended, and attributions of importance are likely to differ between different social and cultural groups. This way, importance information can assist in removing redundant issues before the scale is finalised and subsequently reduces the burden of

<sup>☆</sup> See Appendix for details.

completion for respondents. Secondly, importance information may have clinical value as a heuristic to guide those who deliver health services. Where particular features of QOL are known to be highly valued then clinicians know that if they can satisfy these needs as a priority, their patients will be more contented with their service.

Some measures do examine the importance of different aspects of QOL. Implicit to procedures used in preference-based outcome measures are patient decisions about the relative importance of different health care situations that might affect their mortality and morbidity [6]. Also the importance of elicited or provided dimensions is assessed by some goal attainment measures [7], and in standardised scales e.g. quality of life index (QLI) [8], subjective quality of life profile (SQLP) [9]. Individualised techniques often direct respondents to identify, prioritise and rate or weight important QOL areas e.g. Schedule for the evaluation of individual quality of life (SEIQOL) [10], patient generated index (PGI) [11] and Repertory Grid technique [12].

Advantages and disadvantages of this approach are illustrated by considering two hypothetically similar patients receiving post-operative hospital care for hernia repair who have been asked to identify those aspects of QOL that are most important to them right now. One cites intense pain as very important, while another, attributing an equal level of importance, may describe the worry of not being able to water his tomato plants during a summer drought. This example shows how people ascribe high importance to different and sometimes not always conventional issues. It also serves to suggest that if valid comparisons need to be made between individuals (rather than a number of times for the same individual) only those issues that are agreed by social consensus to be highly important and relevant should be offered for assessment. However to focus on importance alone may not in itself provide a good assessment of QOL. It is argued here that the importance of QOL may be necessary but not sufficient, to provide an overall assessment of the concept. Ratings of how satisfied or bothered people are with their life may need to be integrated with an assessment of importance.

The value of the importance of QOL is reflected in some definitions although is not always explicitly stated. Campbell et al. [13] proposed that

QOL is about the perceived discrepancy between real and ideal states suggesting that QOL would be higher when this discrepancy was small, and lower if it was pronounced [13]. Similarly, Calman says 'A good QOL can be said to be present when the hopes of an individual are matched and fulfilled by experience. The opposite is also true: a poor QOL occurs when the hopes do not meet with the experience' [14]. Cella & Tulsky [15] defined QOL as 'the importance of people's subjective perceptions of their current ability to function, as compared with their own internalised standards of what is possible or ideal'. These definitions have affinity with the WHOQOL Group definition of QOL as 'an individual's perception of their position in life, in the context of the culture and value systems in which they live, and in relation to their goals, expectations, standards and concerns' [16]. Although the importance of different features of quality is not explicitly stated, it is implied, and hence provides theoretical guidance for the present work. A model of 'needs' is also discussed by Hunt & Leplege [17]. Within these definitions, comparison processes are cited as the common vehicle through which judgements about QOL are made [18] but to date, this position statement has not been subjected to a direct empirical test.

The WHOQOL was developed and standardised to measure QOL cross-culturally. It enables self-reported QOL to be measured in relation to health and is a generic instrument for use with many disorders and well people. Covering 25 dimensions (facets) of QOL, it is now available in around 50 languages. The WHOQOL-100 core items were derived following a program of qualitative and quantitative work agreed by an international research collaboration and initially carried out simultaneously in 15 centres world-wide [4, 19, 20]. Because different cultures would be likely to ascribe a distinctive profile of importance values to QOL dimensions, WHOQOL importance items were designed to test this. Their results provided useful evidence for facet selection during instrument development [20]. They were originally intended to weight the core scores with the aim of improving cross-national equivalence between language versions [21] but during development, new research showed that the products resulting from a multiplication of core and importance scores would mislead interpretation, because this

procedure had the potential to produce scores that could mask the fact that a number of people all receiving the same score had quite different experiences of QOL and health [22]. Hence the same score would not enable a valid discrimination to be made between these psychologically different individuals. So the WHOQOL group did not use the importance ratings in scoring but the data remains of intrinsic practical and theoretical interest.

Taking the WHOQOL definition as a starting point, we assumed that the importance that people attached to various aspects of their QOL is influenced by their '...goals, expectations, standards and concerns' [16] and *vice versa*. The four options implied by this definition are operationalised and exemplified as follows: (1) A person with perceived poor mobility and who also believes that it is very important to be mobile may have poor QOL because the differential between these two perspectives is large. As their 'goals, expectations, standards and concerns' are unmet, there is likely to be dissatisfaction or distress. (2) Someone with perceived poor mobility but who see their mobility as relatively unimportant will also report poor QOL, but their QOL may less poor than in (1) because lower aspirations are fulfilled and mobility of less concern. (3) Where mobility is highly important and the person feels very mobile there should be a good QOL because high expectations are met and as with (2), there is little disparity. (4) Where mobility is perceived as good but unimportant, despite the large differential this is unlikely to detract from QOL because of its lack of importance, indeed expectations may be exceeded. This study aims to investigate whether those who rate their QOL as poor and see these aspects as highly important have the very poorest QOL compared to others, because this large differential represents unfulfilled goals, expectations, standards and concerns. A large survey of health-related QOL world-wide, enabled an empirical test of this hypothesis.

## Method

### Design

The cross-sectional data for this study was collected simultaneously in 15 field centres which had been selected to provide international contrasts in

terms of health service provision, industrialisation, geographical region, and other important indicators of QOL like the dominant religion, perceptions of self and time, and the role of the family [20]. They were located in Melbourne, Australia; Zagreb, Croatia; Harare, Zimbabwe; St Petersburg, Russia; Bangkok, Thailand; Paris, France; Bath, England; Panama City, Panama; Madras, India; Barcelona, Spain; Tokyo, Japan; Tilburg, the Netherlands; Seattle, USA; New Delhi, India and Beer Sheeva, Israel. Using a commonly agreed international protocol, each centre obtained questionnaires from a minimum of 300 participants using a quota sample, where a minimum of 50 would be well and 250 sick. They recruited a heterogeneous sample of patients from a wide range of conditions and disorders. A  $2 \times 2$  design was used to target equal numbers of gender and age groups (age was bisected at 45 years). All were adult, in the terms defined by each culture.

### Measures

#### *The WHOQOL-100*

The WHOQOL-100 was derived from a pilot version containing 235 international core items and 41 importance items. Five-point Likert ratings were developed using a visual analogue scale methodology to obtain interval measurement [21]. Only importance items were rated from 'not at all important' (1) to 'extremely important' (5). An example of a core item is 'How safe do you feel in your daily life?' and a compatible importance item is 'How important to you is feeling physically safe and secure?' In a few cases there were necessary linguistic similarities between the wording of core and importance items but most importance items were quite different in style and content.

The 100 core items of the WHOQOL-100 are organised into 25 facets of QOL, which are grouped hierarchically into six domains with 32 importance items appended [20] (see Table 1). This extracted set of 132 items form the basis of the analyses presented in this paper. A General facet (G) assessed overall health and general QOL. During the development of the WHOQOL-100, importance ratings were included in the selection of the final 25 facets, so every facet had high international consensus, being important to very important with a mean international rating

**Table 1.** Pearson correlations between importance (I) scores with corresponding WHOQOL facet means (core), the General Facet (G) and between C-I difference scores with G

	Core (C)	General (G)	Difference C-I vs. G
Domain 1: Physical health			
F1 Pain & discomfort	-0.17*	-0.02	0.33*
F2 Energy & fatigue	0.01	0.03	0.46*
F3 Sleep & rest	0.03	-0.01	0.42*
Domain 2: Psychological			
F4 Positive feelings	0.13*	0.09	0.44*
F5 Thinking, learning, memory etc	0.18*	0.10	0.31*
F6 Self-esteem	0.21*	0.12*	0.36*
F7 Body image & appearance	0.01	0.13*	0.17*
F8 Negative feelings	-0.15*	-0.03	0.41*
Domain 3: Levels of independence			
F9 Mobility	0.07*	0.02	0.37*
F10 Activities of daily living	0.06*	0.08*	0.44*
F11 Dependence on medication	0.14*	0.09*	0.27*
F12 Working capacity	0.25*	0.09*	0.43*
Domain 4: Social relationships			
F13 Personal relations	0.20*	0.16*	0.31*
F14 Practical social support	0.28*	0.03	0.32*
F15 Sex	0.27*	0.11*	0.20*
Domain 5: Environment			
F16 Physical safety & security	-0.01	0.05*	0.27*
F17 Home environment	0.22*	0.21*	0.25*
F18 Financial resources	-0.17*	-0.05*	0.40*
F19 Health & social care	0.07*	-0.01	0.34*
F20 Information & skills	0.22*	0.14*	0.27*
F21 Recreation & leisure	0.28*	0.14*	0.35*
F22 Physical environment	-0.05*	0.15*	0.10
F23 Transport	0.02	0.04*	0.25*
Domain 6: Spirituality, religion & personal beliefs			
F24 Spirituality	0.57*	0.14*	0.13*

\* Significant at  $p \leq 0.001$ .

between 3.0 and 5.0. Some facets had more than one importance item where there had been any ambiguity about which aspect of the concept might best represent importance for that facet e.g. separate items on happiness and contentment for positive feelings. The time-frame for core items enquired about the last 2 weeks but no time-frame was specified for the WHOQOL importance items. Items, response scales and instructions were translated and back-translated in the centres according to an agreed WHO methodology and translations were approved centrally [23]. Although some centres developed national core and importance items appropriate and specific to their language and culture, these are not included in the present analysis [24].

The WHOQOL-100 shows good internal consistency reliability: Cronbach's  $\alpha$  is above 0.7 for almost all facets (0.65–0.93); for General QOL (G) it was 0.84, and for domains 0.71 to 0.86. Test-retest reliability over 2–8 weeks ranged from 0.68 to 0.95, and sick-well discriminant validity was good for facets and domains. The construct and structure of the WHOQOL-100 has been validated using exploratory factor analysis [19, 20].

#### Procedure

All participants completed the pilot version of the WHOQOL-100 followed by the importance items at the same session. Socio-demographic questions on age, sex, educational level, marital status, oc-

cupation and health status included as standard in the WHOQOL-100, were also completed.

### *Analysis*

Statistical analysis was carried out using computing software SPSS Version 10. Before completing these analyses, the single 'best' importance item was selected to reduce the items to one per facet, so that there would be equivalence between all facet importance scores.<sup>1</sup> Frequency distributions showed that ratings were mostly 3, 4 or 5 on the five-point importance scale, due to the prior selection of the most important facets for the WHOQOL-100, so skewed scores were log transformed to normalise distributions.

Correlational analyses at Stage 1 examined the association between the core and importance items of the WHOQOL-100 and also with general (G) QOL. In view of their common QOL origins, a modest but significant positive correlation (Pearson  $r$ ) between core and importance was expected for all facets. It was expected that the relationship would not be strong, as those with poor QOL might downplay importance as a coping strategy, and those with the best QOL upgrade its importance. General QOL ratings were used as a benchmark against which to compare the contribution of the importance items with the core, as G is not an integral part of the WHOQOL-100 scoring. The large number of cases and of correlations demanded that a more conservative  $p$ -value ( $p \leq 0.001$ ) was applied to all results.

<sup>1</sup>Four facets in the WHOQOL-100 contained more than one importance item: positive feelings (F4), thinking etc (F5), access to health care (F19) and information & skills (F20). Tests were conducted to identify the single best performing item that would be representative of the parent facet. Ten items were examined following psychometric selection procedures used by the WHOQOL Group [18, 19]. Item means and standard deviations were all acceptable, ranging from 3.7 to 4.2 (SD 0.8–1.0). Correlations between items in the four key facets with other importance items within the relevant domain were significant (0.29–0.72) and intra-facet items also correlated well (0.54–0.72). Individual items within F19 and F20 showed consistent weaknesses in performance across several tests and were removed to leave the best item in each case. Hierarchical multiple regression analysis (using the general facet importance score as dependent variable) enabled selection of an item for F4 and F5.

Differences between the core and importance ratings were expected to be a useful vehicle for understanding and interpreting the WHOQOL definition. Differences between the mean facet core score and its respective facet importance score were calculated, and ranged on a continuum from  $-4$  where core scores were low (poor QOL) but importance was high, to  $+4$  where QOL was scored as good but importance was low. These differences were correlated with G to examine their relationship to general health and overall QOL. If a large negative difference in scores represents poorest QOL as predicted, e.g. a core score of 1 and importance scoring 5, then these scores would be expected to be closely associated with poor overall QOL (low scores). Conversely, a small positive or negative difference e.g. core = 4, importance = 5, would be expected to indicate better QOL, and would be closely associated with higher G scores, so a significant positive correlation was predicted for each facet between difference scores and G.

To examine whether the inclusion of the importance ratings provides more information than the core scores alone, first the correlation between the core and general facets was investigated then the association was retested controlling for the impact of importance, using partial correlation. If the relationship between the core and general QOL is significantly reduced when importance is removed, then importance makes a substantial contribution to the assessment of perceived QOL. Lastly, stepwise multiple regression was carried out on the 24 facet differences as independent variables (core minus importance) to investigate the extent to which they predicted overall QOL. The dependent variable was the mean of the general QOL core items. The aim was to find out which of these facet differences might have the best predictive value. The  $R^2$  adjusted and change values were inspected with  $F$ -values.

A more demanding test of the hypothesis would be to compare the difference scores from subsamples allocated to groups based on their reported level of QOL – poorest, poor, good and best QOL – and this was carried out at Stage 2. Histograms confirmed skewed frequency distributions with few individuals rating importance as 1 or 2 on the five-point scale as expected, so a score of  $<4.0$  was designated as lower importance in this analysis. The normally distributed core scores

could be bisected, so those scoring 3.0 or less were designated as having a poorer QOL. Together these two scoring criteria provided an algorithm for uniquely allocating each participant to one of four groups, and therefore level of QOL: Group 1 – poorest QOL (core  $<3$  and importance  $>4$ ), Group 2 – poor QOL (core  $<3$  and importance  $<4$ ), Group 3 – good QOL (core  $>3$  and importance  $<4$ ), Group 4 – best QOL (core  $>3$  and importance  $>4$ ). It was predicted that Group 1 with the largest negative differential, would report the poorest QOL of any group and Group 2 with poor QOL, would show a negligible differential arising from low core and low importance scores. Those with good QOL in Group 3 would show a positive differential between core and importance scores and Group 4 would be likely to have the best QOL, as reflected by the highest core and importance scores and a negligible difference. The two groups with poor and poorest QOL were expected to be distinguishable from each other on the basis of the size of the core-importance differences, as those with the poorest QOL would have larger negative differences. This prediction was tested using analysis of variance with paired comparison tests (Scheffe).

Finally, having looked at the facets, we asked whether the differential could discriminate between different levels of overall QOL. Mean difference scores were calculated across all 24 facets (not G) for individuals, and these values were entered into a one-way ANOVA with *post hoc* comparisons (Scheffe test). The dependent variable was a five-point General QOL item: ‘How would you rate your QOL?’ This way, the mean difference scores could be compared for people who rated overall QOL as very poor (1), with those who rated it higher using points 2–5 on the scale. It was predicted that those with the poorest QOL would show the biggest negative differences, and a decrease in the size of differences was predicted for those with other levels of QOL moving across the continuum from 1 to 5.

## Results

### *Sample description*

A total of 4802 participants in 15 centres completed the WHOQOL-100 with its importance

items. The centre sample size ranged from 286 to 412, and the mean age from 38 to 48 years ( $SD = 14$ – $20$ ). Women constituted 49–63%, and sick respondents between 70 and 84% of the samples from each centre. The sample is further described elsewhere [20].

### *Stage 1: Investigation of the relationship between the core and importance items*

The results in Table 1 showed that although 18 of the 24 correlations between core and importance ratings (core) were highly significant, the associations were very small in every facet. This consistent lack of associations is particularly noteworthy in view of the very large numbers in the sample. The exception to this trend was the facet/domain of spirituality, religion and personal beliefs ( $r = 0.57$ ). Furthermore, negative correlations occurred in five facets although a positive association was predicted.

Correlations between importance items and the G core means showed weaker correlations (see Table 1 – G) and while 15 out of 24 correlations were significant, the strongest was only  $r = 0.21$ . As predicted from theory, there is little consistent linear association between the importance of QOL and QOL as measured by the corresponding core or general facets. One interpretation of these weak but consistent findings is that the information provided by the two types of items about QOL is relatively dissimilar. Despite their common conceptual origins, they appear to assess different perspectives on QOL, and in a relatively independent way. This suggests that information derived from the importance items might be used to supplement that of the core in the process of making a better evaluation of QOL.

When core-importance differences were correlated with the General facet (Table 1) the hypothesis was supported, as all correlations were found to be highly significant ( $p < 0.001$  or  $p \leq 0.01$ ) and showed a much stronger set of associations than in previous analyses. This provides limited support for the view that assessing the difference between core and importance scores rather than just inspecting low core scores might provide a better means of identifying those with the poorest QOL.

**Table 2.** To show correlations between core and general quality of life ratings and the impact of importance on these, using partial correlation

WHOQOL Facet	General vs. core	General vs. core controlled for importance
Pain	0.48	0.43
Energy	0.59	0.55
Sleep	0.58	0.55
Positive feelings	0.65	0.60
Thinking etc	0.50	0.43
Self-esteem	0.56	0.53
Body image	0.38	0.36
Negative feelings	0.57	0.53
Mobility	0.45	0.40
Activities	0.61	0.57
Dependence	0.41	0.37
Work	0.54	0.50
Relations	0.58	0.55
Support	0.42	0.39
Sex	0.41	0.40
Safety	0.42	0.36
Home environment	0.47	0.43
Finance	0.49	0.44
Health care	0.43	0.35
Information	0.48	0.41
Leisure	0.57	0.53
Physical environment	0.32	0.31
Transport	0.36	0.29
Spirituality	0.27	0.24

Table 2 shows that for every facet, correlations between the core and general facets were smaller when the impact of importance on these facet scores was controlled for (using partial correlation), than in the direct correlation between the two. Despite the relative similarity in the size of these two sets of correlations, the consistent reduction in size of the correlation for every facet when importance was controlled for, was confirmed as a highly significant difference, using a paired-samples *t*-test of the two sets of correlations shown in Table 2 ( $t = 11.19$  ( $df = 22$ )  $p < 0.0001$ ). It is worth noting that the facets with the largest correlations tend to represent dimensions that are commonly found in many multi-dimensional profiles of health-related QOL. However the correlation sizes of G vs. core scores in Table 2 are considerably greater than correlations in Table 1 for the difference between core and importance with G. This consistently lower level of association where importance is included suggests that the addition of importance information may in fact

detract from a fuller assessment of QOL, and hence the result goes contrary to our hypothesis.

Multiple regression analysis showed that when differences for all 24 facets were entered, 18 facets predicted 39.5% of the variance in overall QOL. The most important of these was positive feelings where 22% of the variance was accounted for, followed by energy and fatigue (7%), financial resources (4%), sleep and rest (2%), and working capacity (2%). The other 13 facets of QOL predicted <1% each. The total variance predicted is not large, with positive feelings accounting for more than half of the effect. This shows that those who believe that having a high level of positive feelings is important to their QOL and who report poor positive feelings indicate that this difference has a substantial effect on their overall QOL. It also provides some support for the view that happiness and contentment as key issues within the positive feelings concept may have universality. However, when the regression was repeated substituting core scores alone for core-importance differences, then the core scores explained a substantially higher percentage of the overall variance at 72%, so our prediction that importance information will improve QOL assessment must be rejected.

#### *Stage 2: Testing the QOL differential*

Table 3 examines the differential of differences after every respondent had been allocated to one of four possible groups using an algorithm to designate them as having the poorest, poor, good or best QOL. Group means for facets are presented in Table 3 with the mean core-importance differences (brackets), and the difference between group differences.

Overall, the *F*-values showed a consistent, highly significant main effect for groups, for every facet. Paired comparisons between groups showed that these differences were greatest for virtually all combinations where the extremes of the QOL continuum were evaluated together i.e. poorest vs. good (3) or best (1); poor vs. good (4) or best (2); that is, 95 out of 96 results ( $24 \times 4$ ) are significant for the above combinations. As expected, the largest negative differences were found for comparisons between poorest and poor with best (1 & 2). The effect was less consistent and smaller where



**Table 3.** Analyses of variance of facets, comparing core-importance differences for different levels of QOL<sup>a</sup>

Facets	<i>F</i> -value	Mean core (and core-importance difference) scores				Comparison of core-importance differences for different pairs of groups					
		Group 1 Poorest QOL	Group 2 Poor QOL	Group 3 Good QOL	Group 4 Best QOL	1. Poorest & Best	2. Poor & Best	3. Poorest & Good	4. Poor & Good	5. Good & Best	6. Poorest & Poor
N		106–589	38–332	130–2132	1987–4368						
Pain & discomfort	288.9*	2.94 (–1.60)	2.98 (–1.56)	3.53 (–1.00)	3.65 (–0.89)	–0.71*	–0.67*	–0.59*	–0.55*	–0.13*	–0.06
Energy & fatigue	789.0*	2.87 (–1.57)	2.86 (–1.56)	3.52 (–0.92)	3.71 (–0.87)	–0.84*	–0.84*	–0.65*	–0.66**	–0.18*	–0.09
Sleep & rest	388.8*	2.53 (–1.72)	2.60 (–1.65)	3.51 (–0.73)	3.51 (–0.73)	–0.97*	–0.90*	–0.97*	–0.91*	–0.15	0.13
Positive feelings	590.9*	2.63 (–1.98)	2.68 (–1.60)	3.44 (–0.84)	3.62 (–0.66)	–1.0*	–0.95*	–0.81*	–0.77*	–0.18*	–0.10
Cognitions	271.9*	2.64 (–1.49)	2.61 (–1.46)	3.31 (–0.82)	3.50 (–0.63)	–0.86*	–0.89*	–0.67*	–0.70*	–0.18*	–0.17
Self-esteem	393.7*	2.73 (–1.43)	2.78 (–1.38)	3.45 (–0.71)	3.60 (–0.56)	–0.86*	–0.81*	–0.72*	–0.67*	–0.14*	–0.10
Body image	188.7*	2.90 (–0.69)	2.80 (–0.79)	3.36 (–0.23)	3.57 (–0.02)	–0.67*	–0.76*	–0.46*	–0.57*	–0.21*	–0.05
Negative feelings	434.9*	2.80 (–1.35)	2.93 (–1.22)	3.50 (–0.65)	3.68 (–0.47)	–0.88*	–0.75*	–0.71*	–0.58*	–0.17*	–0.12***
Mobility	263.7*	2.80 (–1.49)	2.95 (–1.34)	3.42 (–0.87)	3.54 (–0.75)	–0.74*	–0.60*	–0.61*	–0.47*	–0.13*	–0.15*
Activities of daily living	506.8*	2.64 (–1.73)	2.66 (–1.69)	3.37 (–1.00)	3.60 (–0.77)	–0.96*	–0.95*	–0.72*	–0.71*	–0.23*	–0.09
Dependence on medication	216.6*	2.97 (–0.98)	2.92 (–1.03)	3.46 (–0.49)	3.59 (–0.36)	–0.62*	–0.66*	–0.49*	–0.53*	–0.13*	–0.13
Work capacity	406.8*	2.74 (–1.44)	2.88 (–1.30)	3.45 (–0.73)	3.61 (–0.57)	–0.87*	–0.73*	–0.70*	–0.57**	–0.16*	–0.14***
Personal relations	414.9*	2.57 (–1.47)	2.50 (–1.53)	3.30 (–0.67)	3.55 (–0.42)	–0.98*	–1.06*	–0.73*	–0.81**	–0.25*	–0.17
Practical social support	235.0*	2.78 (–1.02)	2.96 (–0.84)	3.52 (–0.28)	3.53 (–0.27)	–0.75*	–0.57*	–0.74*	–0.56*	–0.10	–0.18*
Sex	192.2*	2.94 (–1.32)	3.01 (–1.32)	3.48 (0.15)	3.60 (0.27)	–0.68*	–0.59*	–0.56*	–0.48*	–0.12*	–0.11
Physical safety & security	268.9*	2.84 (–1.67)	2.84 (–1.67)	3.40 (–0.23)	3.57 (–0.40)	–0.72*	–0.73*	–0.55*	–0.56*	–0.17*	–0.12
Home environment	298.1*	2.80 (–1.28)	2.70 (–1.18)	3.30 (–0.79)	3.60 (–0.49)	–0.75*	–0.86*	–0.46*	–0.56*	–0.30*	–0.10
Financial resources	306.2*	2.90 (–1.25)	3.20 (–0.93)	3.57 (–0.58)	3.71 (–0.44)	–0.74*	–0.55*	–0.61*	–0.41*	–0.14*	–0.20*
Health & social care	241.0*	2.73 (–1.59)	2.89 (–1.43)	3.43 (–0.89)	3.50 (–0.87)	–0.76*	–0.61*	–0.70*	–0.54*	–0.15	–0.17
Information & skills	291.3*	2.73 (–1.06)	2.70 (–1.03)	3.36 (–0.43)	3.55 (–0.24)	–0.81*	–0.84*	–0.63*	–0.66*	–0.18*	–0.07
Recreation & leisure	447.2*	2.89 (–0.97)	2.80 (–0.88)	3.60 (–0.26)	3.69 (–0.11)	–0.80*	–0.85*	–0.71*	–0.76*	–0.12	–0.10
Physical environment	142.6*	3.09 (–0.68)	2.94 (–0.83)	2.83 (–0.94)	3.52 (–0.25)	–0.43*	–0.56*	0.25**	0.22	–0.69*	0.15*
Transport	145.6*	2.93 (–1.05)	3.02 (–0.96)	3.39 (–0.59)	3.52 (–0.46)	–0.58*	–0.50*	–0.45*	–0.37*	–0.13***	–0.07
Spirituality, religion etc	89.5*	3.05 (–0.76)	3.06 (–0.73)	3.46 (–0.35)	3.47 (–0.34)	–0.42*	–0.41*	–0.40*	–0.40*	–0.14	–0.18

Key: \*  $p < 0.001$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.05$ .<sup>a</sup> *F* represents the overall differences between the four groups (poorest QOL, poor QOL, good QOL and best QOL). Ns in columns three to six represent the ranges in terms of the maximum and minimum number of people who provided data for each facet. The four left-hand columns of means also display core minus importance differences in brackets.

more similar groups were compared; for best with good QOL (5), 19 out of 24 facets were significant. But when those with poor and the poorest QOL (6) were compared, only 6 out of 24 facets were significant, and of these, the direction of the difference for physical environment (positive) was against prediction, so the specific hypothesised effect can only be confirmed for five facets. Facets on mobility, practical social support and financial resources showed strong support for the prediction ( $p < 0.001$ ); negative feelings and working capacity were significant but the evidence was weaker ( $p < 0.05$ ). The means and core-importance differences (left hand columns) display the predicted gradient for the majority of facets (18 out of 24) but inconsistent results are seen for body image, dependence on medication, personal relations, sex, safety and security, and physical environment.

The final ANOVA using mean difference scores for accumulated facets showed a large significant main effect confirming discrimination between different levels of QOL as indicated by a five-point rating on a single general item 'How would you rate your QOL'? *Post hoc* comparisons indicated significant differences between every pair of intervals across the five-point scale ( $p < 0.0001$ ) showing that the core-importance differences change across the continuum of rated QOL. Results for each pair of comparisons confirmed that those reporting very poor QOL had the largest mean difference ( $-1.5$ ) ( $n = 83$ ) and those with the best QOL the smallest difference ( $-0.17$ ) ( $n = 578$ ). This pattern is consistent for all the points along the scale with differences becoming incrementally smaller as QOL improves (poor  $-1.1$  ( $n = 434$ ), neither poor nor good  $-0.64$  ( $n = 1270$ ), good  $-0.32$  ( $n = 1756$ )). These results support the overall predicted effect.

## Discussion

In this study we examined the relationship of importance ratings of different dimensions of QOL–QOL assessment. The facets of QOL included in the WHOQOL-100 were known to be important across a large number of countries worldwide [19], but individuals have different priorities about the importance that they give to different dimensions of QOL, and what these importance ratings mean

for individuals is also of interest in this study. We began with the observation that core and importance ratings of QOL showed remarkably little association with each other in correlations, given the size of the sample and the fact that they had been derived from a common QOL stem. But a substantial association would not have been theorised from the WHOQOL Group's definition of QOL which indicates that judgements about QOL are influenced by peoples 'goals, expectations, standards and concerns'. Implicit in this is the importance that they give to these dimensions. We drew from this definition in predicting that a large and negative differential between core and importance ratings would be closely associated with the poorest QOL. Integral to this discussion was a practical question about whether the inclusion of importance items in this assessment, with its attendant burden for the patient, could substantially improve the prediction of QOL.

A series of correlations showed that although there was some initial support for the hypothesis, when compared with core scores alone, the use of differences between core and importance items substantially reduced rather than increased the capacity of the WHOQOL to predict overall QOL and so in rejecting this hypothesis, a general case for including all importance items in assessment could not be supported. However, when respondents were uniquely categorised into four distinct QOL groups on a continuum from poorest QOL to best QOL and the differences tested across the different groups, there was strong support for a main effect across groups and this was reflected in the means of 18 out of 24 facets. However an investigation of comparisons between pairs of groups showed only limited support from 5 out of 24 facets for the more specific hypothesis that those who attach considerable importance to particular aspects of their life and at the same time rate these aspects as very poor (poorest group), have poorer QOL than those who rate it as poor but do not attach much importance to it (poor group). In the few facets where this differential occurs, it shows that aspirations for a good QOL are not being met in terms of people goals, expectations, standards and concerns, and this detracts from a good QOL.

While the results point to the general conclusion that completion of all the additional importance

items would substantially increase the burden to respondents and administrators with no gain, there would be little additional burden from completing these five identified importance items. To identify someone with the poorest QOL it may help to look at the importance given to these five key facets. For these indicators, the poorest QOL appears to be more acutely experienced because being mobile, having sufficient practical social support and financial resources to meet ones needs, being free from negative feelings like depression and anxiety, and having the capacity to work, are all seen as important. These facet difference scores distinguish those with the poorest QOL from those with poor QOL but furthermore the regression showed that the differentials for financial resources and working capacity made significant improvements to the prediction of overall QOL. Together these results show that importance ratings do contribute to the full assessment of certain facets so it is possible that the importance of different selections of facets from this generic profile may be highly salient for specific populations with particular health conditions e.g. HIV infection, rheumatoid arthritis, and in specific contexts like palliative care.

It is also plausible that some dimensions may not be recognised as important until certain events or health problems occur e.g. pain, sleep disturbance, whereas other dimensions like home environment may be consistently important, irrespective of health state, and the presence (or absence) of a problem. This stability issue may begin to explain why the hypothesis is only partially confirmed for a few facets. The results might also be explained by the phenomenon of response shift [25]. There is growing evidence that people alter the meaning of their own self-evaluation as a result of changes in values, internal standards or with a reconceptualisation of QOL and this often occurs in response to a catalyst like a major life event. It is therefore questionable whether those who have adapted to their poor QOL and are resigned to it, really have a better QOL than those whose QOL is poor but who are less resigned to this situation, and this model would begin to explain the similarity of findings between the poorest and poor groups. Because of the emphasis on goal reordering, changing values and standards, and social comparisons as important processes within the response shift model,

there are conceptual commonalities with the definition of QOL as outlined by the WHOQOL group.

At a practical level the differential where it exists, may serve to earmark those who are most likely to need help in restoring a good QOL. However it would be unwise to conclude that people with the largest differential should necessarily have priority health care over those with lower expectations. They may be most distressed by life but they may also be the most powerful and vocal members of society by virtue of education, opportunity, social class, wealth etc that may have given them higher aspirations. Using the differentials in the absence of other indicators or independent data has the potential to lead to undue neglect among those with low aspirations, who might already be the most disadvantaged or marginalised members of society, and for these reasons report less importance. Assessments of 'need' and/or standard of living to stand alongside the two measures from the WHOQOL-100 might be a solution. Caution needs to be exercised that any such procedures do not serve to accentuate existing social inequalities. This problem is however commonly shared with most self-report measures in this field and is not unique to the use of the importance differential recommended here.

Some methodological observations may cast light on the partially confirmed results. Despite the size of the total sample, the smaller numbers of respondents for the testing of some facets in the poorest and poor sub-groups may explain the weak results. Relating to this, although the algorithm readily distinguishes those with poorest QOL from those with good and the best QOL, for the majority of facets it was relatively difficult to distinguish those with poor from the poorest QOL. Secondly, although general core scores were used as an independent index against which to judge this phenomenon, this variable was necessarily already constructed in a style similar to the core WHOQOL items and as such, was only a partially independent benchmark from the data contained in both measures we were attempting to assess. In future, an overall QOL item, perhaps drawn from a different QOL measure, would be a better independent yardstick against which to test the differences hypothesis. Thirdly, the curtailed interval importance scale of effectively three points and consequently few respondents who said that any of

the facets were not important, meant that we were unable to use the full and more sensitive range of five-point interval ratings that could otherwise have been available for analysis. Future work might seek to include some new facets that are known to be relatively unimportant to QOL, for the purpose of a better test of the hypothesis.

The method proposed here could be used with patients at the time of consultation and has the potential to identify QOL areas of risk. A simple way of understanding these differences at an individual level for clinical use might be to graph the two types of scores, which would be a faster method of appraisal than engaging in the laborious calculation of differences. At the same time this information may also be a vehicle for improving communications between clinicians and their patients and to initiate self-management of the 'expert' patient.

### Acknowledgements

The authors wish to express their gratitude to several anonymous referees for insightful comments about their findings.

### Appendix

The WHOQOL Group comprises a co-ordinating group of collaborating investigators in each of the field sites and a panel of consultants. At the World Health Organisation, Dr D. Rex Billington directed the project that was initiated by Dr John Orley and Dr Norman Sartorius. The work reported here was carried out by investigators in 15 field sites: Prof H. Herrman, Dr H. Schofield and Ms B. Murphy, St. Vincent's Hospital, Australia; Prof S. Szabo, Prof Z. Metelko and Mrs M. Piber-nik-Okanovic, University of Zagreb, Croatia; Dr N. Quemada and Dr A. Caria, INSERM, Paris, France; Dr S. Kumar, Chennai Medical College, Madras, India; Dr S. Saxena and Dr K. Chandiramani, All India Institute of Medical Sciences, New Delhi, India; Dr M. Amir and Dr D. Bar-On, Ben-Gurion University of the Negev, Beer-Sheeva, Israel; Dr M. Tazaki, Science University of Tokyo and Dr Aiko Noji, St Lukes College of Nursing Japan; Dr G. van Heck and Dr J. de Vries, Tilburg University, The Netherlands; Prof J. Arroyo Sucre and Prof Picard-Ami, University of Panama, Panama; Prof M. Kabanov, Dr A. Lomachenkov and Dr G. Burkovsky, V.M. Bekhterev Psychoneurological Research Institute, St Petersburg, Russian Federation; Dr R. Lucas Carrasco, University of Barcelona, Spain; Dr Yooth Bodharamik and Mr K. Meesapaya, Institute of Mental Health, Bangkok, Thailand; Prof S.

Skevington, University of Bath, Bath, United Kingdom; Prof D. Patrick, Ms M. Martin and Dr D Buesching, University of Washington, Seattle, United States; Prof W. Acuda and Dr J. Mutambirwa, University of Zimbabwe, Harare, Zimbabwe.

### References

1. The WHOQOL Group. The development of the World Health Organisation Quality of Life Assessment Instrument (the WHOQOL). In: Orley J, Kuyken W (eds), *Quality of Life Assessment: International Perspectives*, Heidelberg: Springer-Verlag, 1994.
2. Fallowfield L. *The Quality of Life: The Missing Measurement in Health Care*. London: Souvenir Press, 1990.
3. Bowling A. *Measuring Disease*. Buckingham, UK: Open University Press, 1995.
4. Skevington SM, Mc Arthur P, Somerset M. Developing items for the WHOQOL: An investigation of contemporary beliefs about quality of life related to health in Britain. *Br J Health Psychol* 1997; 2: 55–72.
5. Kind P. The Euroqol Instrument: An index of health related quality of life. In: Spilker B (ed), *Quality of life and Pharmacoeconomics in Clinical Trials*, USA: Lippincott: Raven Publishers, 1996; 91.
6. Kaplan RM, Feeny D, Revicki D. Methods for assessing relative importance in preference based outcome measures. In: Joyce CRB, O'Boyle CA, McGee H (eds), *Individual Quality of Life approaches to Conceptualisation and Assessment*, Amsterdam: Harwood Academic Press, 1999.
7. Browne J. Selected methods for assessing individual quality of life. In: Joyce CRB, O'Boyle CA, McGee H (eds), *Individual Quality of Life Approaches to Conceptualisation and Assessment*, Amsterdam: Harwood Academic Press, 1999.
8. Ferrans CE, Powers MJ. Psychometric assessment of the quality of life index. *Res Nurs Health* 1992; 15: 29–38.
9. Dazord A, Gerin P, Boissel JP. Subjective quality of life assessment in therapeutic trials: Presentation of a new instrument in France (SQLP: subjective quality of life profile) and first results. In: Orley J, Kuyken W (eds), *Quality of Life Assessment: International Perspectives*, Heidelberg: Springer-Verlag, 1994; 185–195.
10. McGee HM, O'Boyle CA, Hickey A, O'Malley K, Joyce CRB. Assessing the quality of life of the individual – the SEIQOL with a healthy and gastroenterology unit population. *Psychol Med* 1991; 21: 749–759.
11. Ruta DA, Garratt AM, Leng M, Russell IT, MacDonald LM. A new approach to the measurement of quality of life – the patient generated index. *Med Care* 1994; 32: 1109–1126.
12. Dunbar GC, Stoker MJ, Hodges TCP, Beaumont G. The development of the SBQOL – a unique scale for measuring quality of life. *Br J Med Econ* 1992; 65–74.
13. Campbell A, Converse PE, Rogers WL. *Quality of American Life Perceptions, Evaluations and Satisfaction*. New York: Russell Sage Foundation, 1976.
14. Calman KC. Quality of life of cancer patients – a hypothesis: *J Med Ethic* 1984; 10: 124–127.
15. Cella D, Tulsky DS. Measuring quality of life today: Methodological aspects. *Oncology* 1990; 5: 29–38.

16. The WHOQOL Group. Study protocol for the World Health Organisation project to develop a quality of life assessment instrument (the WHOQOL). *Qual Life Res* 1993; 2: 153–159.
17. Leplege A, Hunt S. The problem of quality of life in medicine. *J Am Med Assoc* 1997; 278: 47–50.
18. Skevington SM. Social comparisons in cross-cultural quality of life assessment. *Int J Mental Health* 1994; 2: 29–47.
19. The WHOQOL Group. The World Health Organisation Quality of Life assessment (WHOQOL): Position paper from the World Health Organisation. *Soc Sci Med* 1995; 41: 1403–1409.
20. The WHOQOL Group. The World Health Organisation quality of life assessment (WHOQOL): Development and general psychometric properties. *Soc Sci Med* 1998; 46: 1569–1585.
21. Szabo S, Orley J, Saxena S. An approach to response scale development for cross cultural questionnaires. *Eur Psychol* 1997; 2: 270–276.
22. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*. UK: Oxford Medical, 1995.
23. Sartorius N, Kuyken W. Translation of health status instruments. In: Orley J, Kuyken W (eds), *Quality of Life Assessment: International Perspectives*. Heidelberg: Springer-Verlag, 1994; 3–18.
24. Skevington SM, Bradshaw J, Saxena S. Selecting national items for the WHOQOL: Conceptual and psychometric considerations. *Soc Sci Med* 1999; 48: 473–487.
25. Schwartz CE, Sprangers MAG (eds), *Adaptation to Changing Health: Response Shift in Quality of Life Research*. Washington: American Psychological Association, 2000.
26. Raphael D, Renwick R, Brown I, Rootman I. Quality of life indicators and health: Current status and emerging conceptions. *Soc Indic Res* 1996; 39: 65–88.

*Address for correspondence:* S.M. Skevington, Department of Psychology, WHO Centre for the study of Quality of Life, University of Bath, Bath BA2 7AY, UK  
 Phone: +44-1225-826830; Fax: +44-1225-826752  
 E-mail: s.m.skevington@bath.ac.uk